AMP webinar DOI & Data Policy

Deb Agarwal, You-Wei Cheah







Webinar Objectives



- Explain the AmeriFlux data citation DOIs
 - What they are
 - How we assign them
 - How to provide the information needed for your DOIs
- Describe upcoming changes to data usage policy
 - New open tier
- Provide options for addressing emerging data citation challenges

Digital Object Identifiers



Unique alphanumeric string assigned by a registration agency

```
10. <DOI registration agency ID> / <publisher provided suffix>
```

DOI ties together metadata of an object with a digital location



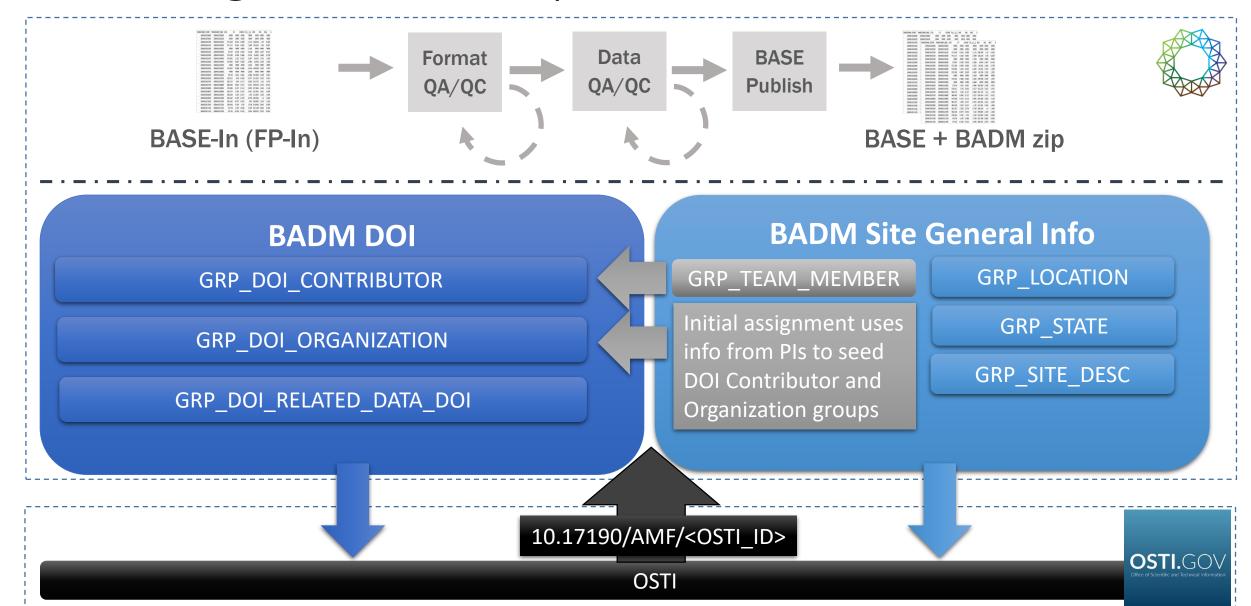
- "Resolvable" Search engines can resolve DOIs
- DOIs cites the data, citing a paper cites just the contents of the paper
- Compliant with FAIR principles: facilitates reproducibility, reusability, interpretation

Current state of AmeriFlux DOIs

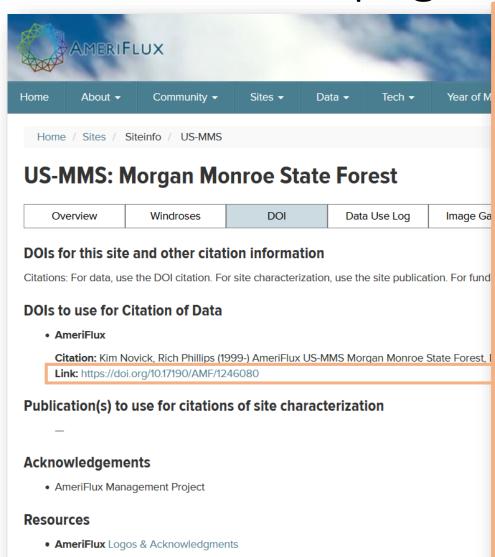


- AMP issues DOIs because sites need to get credit!
- DOIs are assigned to sites with published data products
 - Only AmeriFlux DOI currently applies to the BASE-BADM data product
 - Assigned through the Office of Scientific and Technical Information (OSTI)
 - DOI format: **10.17190/AMF/<Numeric ID>**
- AMP mints DOIs within a month of first data release
- Periodic syncing of DOI metadata with OSTI (about 3-4 times a year)
- Populated using BADM information

DOI assignment and update



AmeriFlux DOI pages



DOI: 10.17190/AMF/1246080

AmeriFlux US-MMS Morgan Monroe State Forest



Dataset Description

This is the AmeriFlux version of the carbon flux data for the site US-MMS Morgan Monroe State Forest. Site Description - Owned by the Indiana Department of Natural Resources (IDNR), the Morgan Monroe State Forest, the site's namesake, is operated thanks to the long-term agreement between Indiana University and IDNR. The first settlers cleared the surrounding ridges for farming, but were largely unsuccessful. The state of Indiana purchased the land in 1929, creating the Morgan Monroe State Forest. Many of the trees in the tower footprint are 60-80 years old, surviving selective logging that ended over the past 10 years. Today, the forest is a secondary successional broadleaf forest within the maple-beech to oak hickory transition zone of the eastern deciduous forest.



Citations

Roman, D. T., Novick, K. A., Brzostek, E. R., Dragoni, D., Rahman, F., Phillips, R. P. (2015), The Role Of Isohydric And Anisohydric Species In Determining Ecosystem-Scale Response To Severe Drought Oecologia, 179(3), 641-654 DOI: 10.1007/s00442-015-3380-9

Other Publications

Sullivan, R. C., Cook, D. R., Ghate, V. P., Kotamarthi, V. R., Feng, Y. (2019), Improved Spatiotemporal



ISEFUL LINKS

People

Opportunities Image Gallery

Events

Logos & Acknowledgments

Tech Blog

Data Blog Safety

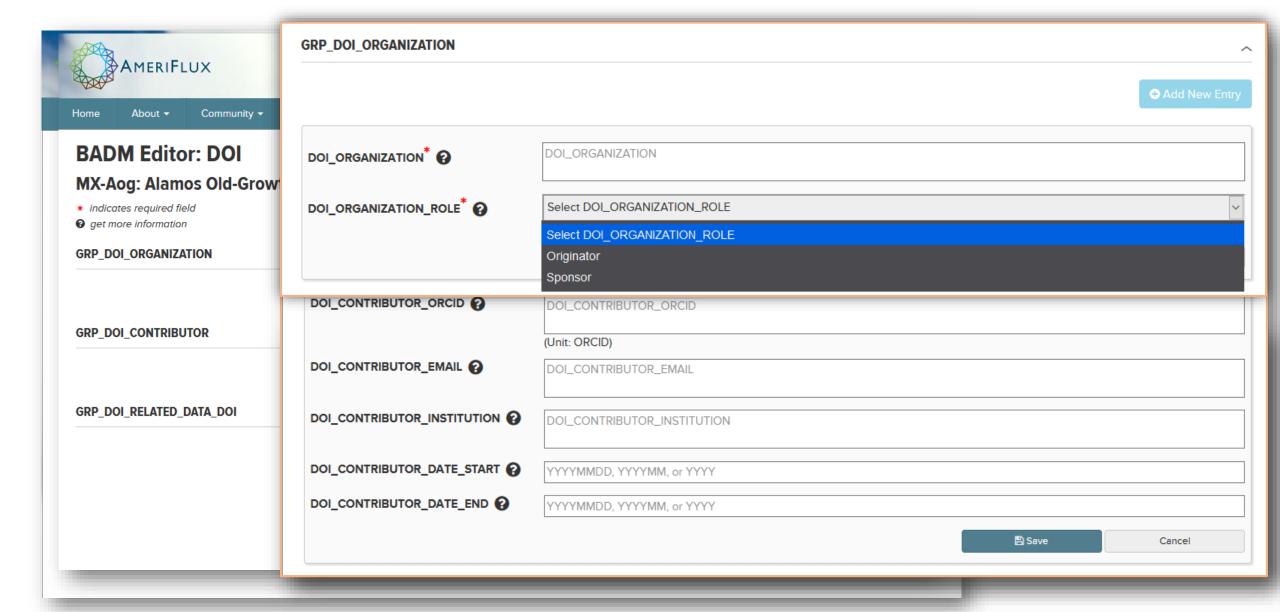
Research Highlights

Publications

Publications

AmeriFlux Flyers

AmeriFlux Online DOI editor



Real-world DOI examples

GRP DOI CONTRIBUTOR

№ 🛍 DOI CONTRIBUTOR NAME Ariane Arias Ortiz DOI CONTRIBUTOR ROLE DOI CONTRIBUTOR ORDINAL

DOI CONTRIBUTOR ORCID 0000-0001-9408-0061

DOI CONTRIBUTOR EMAIL aariasortiz@berkeley.edu

DOI CONTRIBUTOR INSTITUTION University of California, Berkeley

DOI CONTRIBUTOR DATE START

DOI CONTRIBUTOR DATE END

DOI CONTRIBUTOR NAME Joseph Verfaillie

DOI CONTRIBUTOR ROLE

DOI CONTRIBUTOR ORDINAL

DOI CONTRIBUTOR ORCID 0000-0002-7009-8942

DOI CONTRIBUTOR EMAIL jverfail@berkeley.edu

DOI CONTRIBUTOR INSTITUTION University of California, Berkeley

DOI CONTRIBUTOR DATE START

DOI CONTRIBUTOR DATE END

DOI CONTRIBUTOR NAME Dennis Baldocchi

DOI CONTRIBUTOR ROLE

DOI CONTRIBUTOR ORDINAL

DOI CONTRIBUTOR ORCID 0000-0003-3496-4919

DOI CONTRIBUTOR EMAIL baldocchi@berkeley.edu

DOI CONTRIBUTOR INSTITUTION University of California, Berkeley

DOI CONTRIBUTOR DATE START

DOI CONTRIBUTOR DATE END

Dataset Description

Add New Entry

AND S

This is the AmeriFlux version of the carbon flux data for the site US-Mvb M Mayberry Wetland site is a 300-acre restored wetland on Sherman Island. property of Mayberry Farms and managed by the California Department o During Summer 2010, the site was restored from a pepperweed and annu a project managed by Bryan Brock (bpbrock@water.ca.gov). A flux tower e and CH4 fluxes was installed on October 14, 2010. At the time of installation begun after extensive reconstruction of the wetland bathymetry conducte small patches of tules remain within the site, the site is a patchwork of dec remaining vegetation. Currently, there is an intention to flood-to-kill the cu and let the wetland plants propagate naturally, so no additional plant mani-

Dataset Information

Originating Research University of California, Berkeley Organization(s)

USA Country

Sponsor Organization(s) California Department of Water Resources

Site Page http://ameriflux.lbl.gov/sites/siteinfo/US-Myb

Jaclyn Hatala Matthes, Cove Sturtevant, Patty Oikawa, Samuel D Citation

Chamberlain, Daphne Szutu, Ariane Arias Ortiz, Joseph Verfaillie. Dennis Baldocchi (2010-) AmeriFlux US-Myb Mayberry Wetland, Dataset.

DOI: 10.17190/AMF/1246139

Dataset Creator(s)

Name Email

Affiliation

Name

Email

Affiliation

Affiliation

Affiliation

Affiliation

Affiliation

Affiliation

Affiliation

AmeriFlux US-Myb Mayberry Wetland

Jaclyn Hatala Matthes

Cove Sturtevant

Patty Oikawa

Daphne Szutu

Ariane Arias Ortiz

Joseph Verfaillie

Dennis Baldocchi

baldocchi@berkelev.edu

jverfail@berkeley.edu

jaclyn.hatala.matthes@gmail.com

University of California, Berkeley

csturtevant@battelleecology.org

University of California, Berkeley

patty.oikawa@csueastbay.edu

Samuel D Chamberlain

University of California, Berkeley

sam.d.chamberlain@gmail.com

University of California, Berkeley

daphneszutu@berkeley.edu

aariasortiz@berkelev.edu

University of California, Berkeley

University of California, Berkeley

University of California, Berkeley

University of California, Berkeley

https://doi.org/10.17190/AMF/1246139

Data Download Compatibility

Download Data Here

DOI CONTRIBUTOR NAME Jaclyn Hatala Matthes

DOI CONTRIBUTOR ROLE

DOI CONTRIBUTOR ORDINAL

DOI CONTRIBUTOR ORCID 0000-0001-8999-8062

DOI CONTRIBUTOR EMAIL

jaclyn.hatala.matthes@gmail.com

DOI CONTRIBUTOR INSTITUTION University of California, Berkeley

DOI CONTRIBUTOR DATE START 20101014

DOI CONTRIBUTOR DATE END 20130513

DOI CONTRIBUTOR NAME Cove Sturtevant

m iii

DOI CONTRIBUTOR ROLE

DOI CONTRIBUTOR ORDINAL

DOI CONTRIBUTOR ORCID 0000-0002-0341-3228

DOI CONTRIBUTOR EMAIL csturtevant@battelleecology.org

DOI CONTRIBUTOR INSTITUTION University of California, Berkeley

DOI CONTRIBUTOR DATE START

DOI CONTRIBUTOR DATE END

DOI CONTRIBUTOR NAME Patty Oikawa

DOI CONTRIBUTOR ROLE

Author

DOI CONTRIBUTOR ORDINAL

DOI_CONTRIBUTOR_ORCID

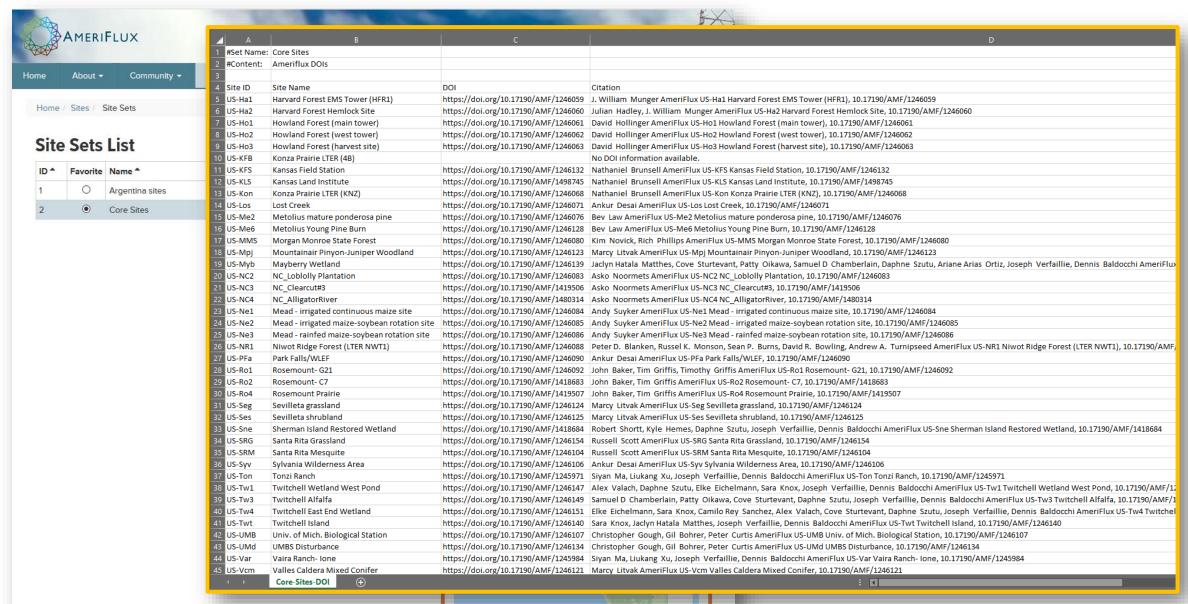
0000-0001-7852-4435 DOI CONTRIBUTOR EMAIL patty.oikawa@csueastbay.edu DOI CONTRIBUTOR INSTITUTION University of California, Berkeley

DOI CONTRIBUTOR DATE START

DOI CONTRIBUTOR DATE END 20161013



DOIs and Site Sets - taming multiple sites' DOIs



DOIs in the context of multiple flux-data products

- Separate DOIs issued for each data product assigned by OSTI via AMP
 - BASE-BADM (AmeriFlux) 10.17190/AMF/<Numeric ID>
 - FLUXNET2015 (FLUXNET) 10.18140/FLX/<Numeric ID>
 - FLUXNET-CH4 (FLUXNET community data product) 10.18140/FLX/<Numeric ID>
- DOI Contributors and Organizations for FLUXNET products are seeded using AmeriFlux DOI information unless explicitly collected for the product
- Changes to DOI Contributors through online editor applies to AmeriFlux data products only
- Different set of contributors between data products is possible (contact AMP to update for now – ameriflux-support@lbl.gov)
- DOI Organizations are shared throughout data products
- DOIs between different products are linked automatically





AmeriFlux Data Usage Policy

AmeriFlux Data Usage Policy Modernization

Current Usage Policy Challenges

- Requirement to contact PI
 - Not compatible with many emerging bulk data usage scenarios
 - Contact is already occurring automatically on download
 - Burden of later contact is high if using more than a few sites
 - Retirements and deaths potentially leave data in limbo
- No explicit redistribution rights
 - No actual right for others to make data available e.g. PEcAn, R Tool, etc.
 - Transfer of AmeriFlux data by AMP to a different data repository (allowed?)
- No explicit adaptation rights

Creative Commons CC BY 4.0



You re re to.

- •Share copy a rid mbu tlematerial in any medium or format
- •Adapt re. x, tall form, and build upon the material for any purpose, even comine ciarly.

Un to the following the terms:

Attribution — You must give appropriete credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

https://creativecommons.org/licenses/by/4.0/

CC BY 4.0 and Similar Policies Becoming Standard

Data Resource	Policy	Usage	Citation	Author Contact Required
FLUXNET2015 paper	CCby4	unrestricted	required	no
ESS-DIVE	CCby4 (and CC0)	unrestricted	required	no
NEON	Custom Similar to CCby4	unrestricted	expected	no
ICOS Carbon Portal	Custom Similar to CCby4	unrestricted	required	no
ARM	Custom Similar to CCby4	unrestricted	required	no
ESGF CMIP	Custom Similar to CCby4	unrestricted	expected	no
Figshare	Creative Commons and other policies	Based on policy	Based on policy	no
LTER <u>Environmental</u> <u>Data Initiative</u>	Specified by data producer	Based on policy	Based on policy	Based on policy
Arctic Data Center	CCby4	unrestricted	required	no
NASA	Mostly CC0	unrestricted	recommended	no

CC BY 4.0 Advantages

- Compatibility with FLUXNET, ICOS, and NEON data policies
- Enables value added tools to be built
- Enables AmeriFlux data to be used in papers submitted to publishers requiring open data
- Enables data papers to be published about AmeriFlux data

• ...

AmeriFlux Data Usage Policy (by end of 2021)

- New AmeriFlux usage policy becomes CCby4 (or similar)
 - Data free for use and reuse
 - Proper citation/acknowledgement still required (specifics specified)
 - Sites receive data download notification
- Support sites who remain on older "AmeriFlux policy"
 - Data quality checks and BASE publish
 - Limited data and tech services e.g. no ONEFlux processing, lower priority for services
- One data policy per site
- AmeriFlux core sites and NEON have indicated that their data can be CCby4

AmeriFlux Data Citation Policy Modernization

Current citation policy challenges

- Citation of a paper about the site does not cite the data
 - Cite the data DOI for citation of the data
 - Cite the site papers for citation of site characterization
- No recognition of other contributors (e.g. AmeriFlux data and tech teams)
- Difficult to cite a large number of sites in reference section
 - AmeriFlux citations end up in supplementary materials (rarely indexed)

Changes to DOI Citation Info

Updated citation guidance

DOIs to use for Citation of Data

AmeriFlux

Citation: Sebastien Biraud, Marc Fischer, Stephen Chan, Margaret Torn (2002-) AmeriFlux US-ARM ARM Southern Great Plains site- Lamont, Dataset. https://doi.org/10.17190/AMF/1246027

Link: https://doi.org/10.17190/AMF/1246027

Publication(s) to use for citations of site characterization

 Fischer, M. L., Billesbach, D. P., Berry, J. A., Riley, W. J., Torn, M. S. (2007) Spatiotemporal Variations In Growing Season Exchanges Of Co2, H2o, And Sensible Heat In Agricultural Fields Of The Southern Great Plains, Earth Interactions, 11(17), 1-21.

Upcoming changes

- Add data version to the citation (e.g. Ver. 17-5)
- Updates to product name (e.g. BASE-BADM)
- Specify publisher (e.g. AmeriFlux)

Collective Citations for AmeriFlux

AmeriFlux Collective Citation – Options

- Implement a collective data citation method
 - Provide a citation that can be included in the reference section
 - Help enable tracking of citations to data
 - Enable credit for other data contributors (e.g. AMP data curators)
 - Three options for when large number of sites used
 - Data papers as data citation
 - Dynamic data citation
 - Data collections as data citation
- Could request table of sites/versions in paper still
- Still receive data download notifications

Data Collections

- Obtain a DOI for an AmeriFlux collection (e.g. AmeriFlux2020)
 - E.g. Includes data through year xxxx for CC BY 4.0 sites
 - Can contain authorship in citation if desired
 - Include papers about data preparation and quality methods
 - Can be cited in the reference section of a paper
 - Landing page for collection and data download provided by AmeriFlux
- Imprecise citation in terms of what sites used (table of sites/versions could address this)

Data Paper

- Provides a description of the dataset and methods used to produce the data
 - DOI of data paper could be used as citation for the data
 - All contributors could be authors (like FLUXNET2015)
- Imprecise citation in terms of what sites used (table of sites/versions could address this)
- Easier to trace citations of the data
- Greater visibility of the data
- Requires CC BY 4.0

Dynamic Data Citations

- Provide a DOI per request or download
 - Covers a specific set of sites and versions of the data used in a paper
 - Can decide authorship in citation
 - DOI landing page provided by AmeriFlux with info of sites/versions used
 - Cited in reference section
- Tracking usage of data in publications requires significant automation
- Difficult to track citation of an individual site's data (relies on AmeriFlux maintaining statistics)

Questions?

Proposed AmeriFlux Embargo Policy for New AmeriFlux Sites

- What is it?
 - Opportunity to get early feedback on data quality to new sites
 - Only for sites who have never published data in AmeriFlux
 - Must be requested by a site (not automatic)
- What is provided?
 - Automated quality check processing of data feedback (format and data)
 - Manual interpretation of automated QA at AMP discretion
- Additional information
 - No ONEFlux processing (gap-filling and partitioning)
 - No Site visits
 - Data may be published by AMP two years after submission